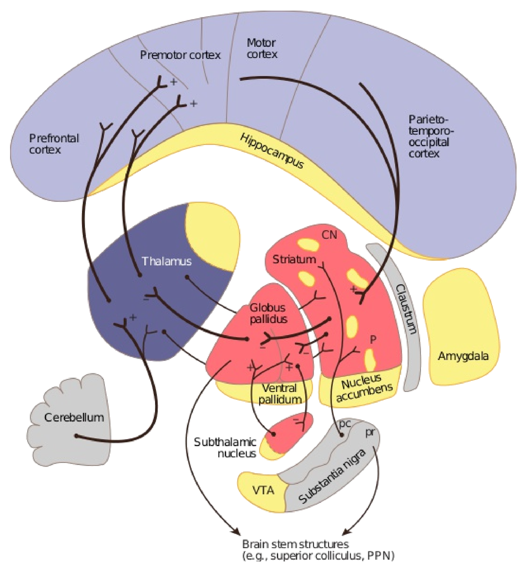
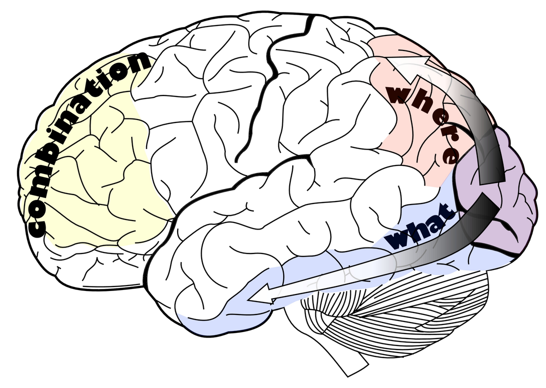
Who's the teacher? Who's the pupil?

Meropi Topalidou, Daisuke Kase, Thomas Boraud, Nicolas P. Rougier

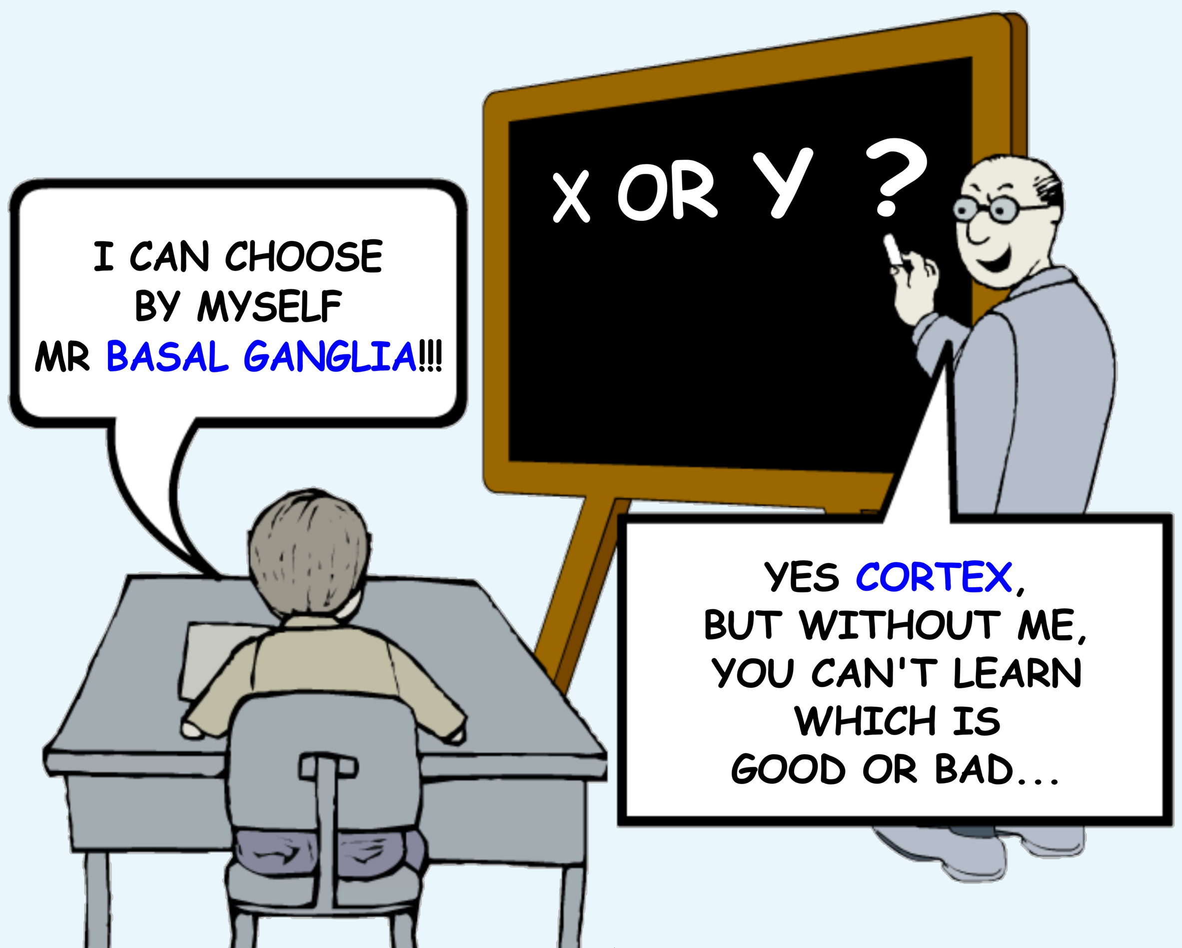


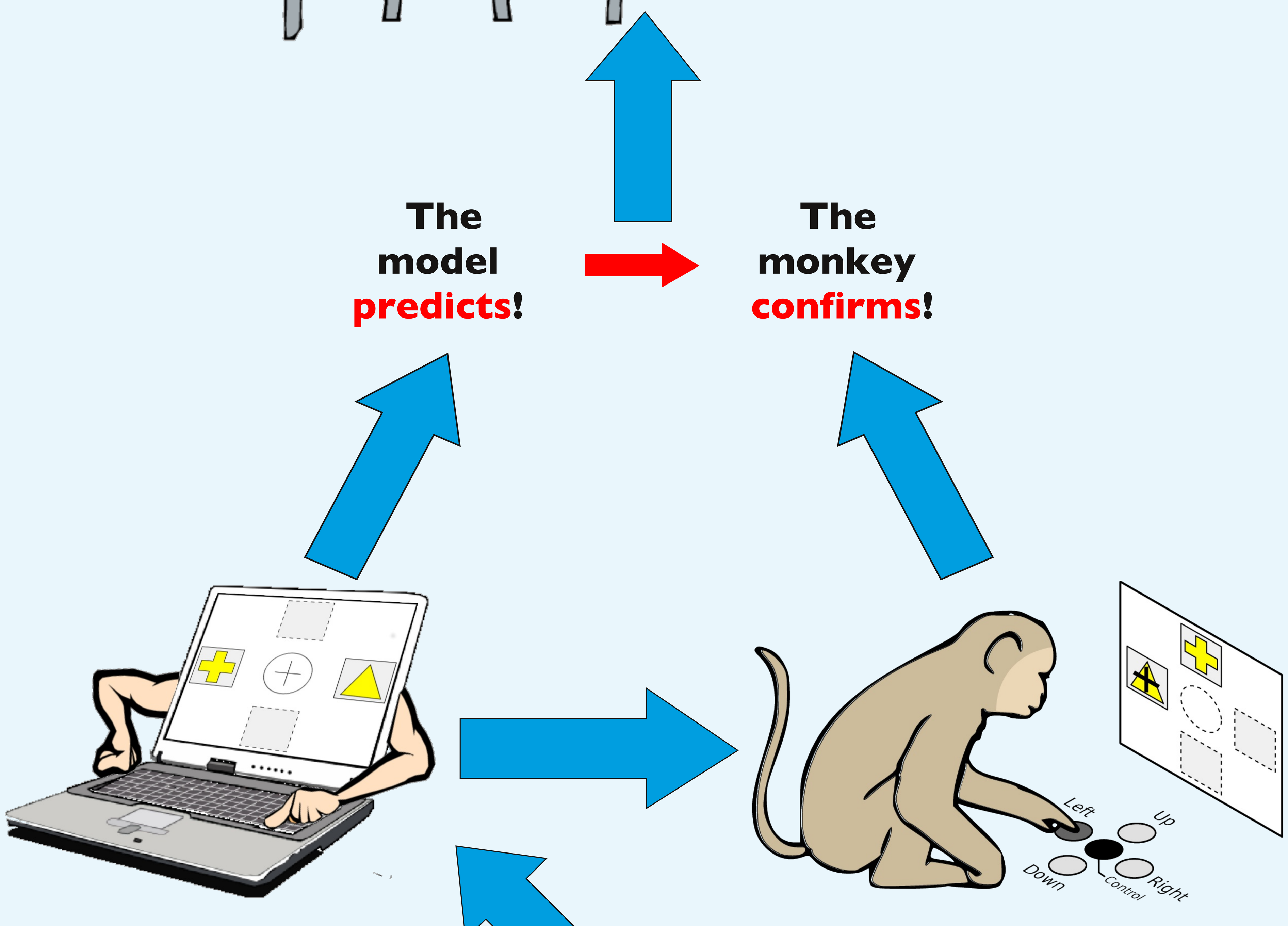
Who's the teacher? Who's the pupil?

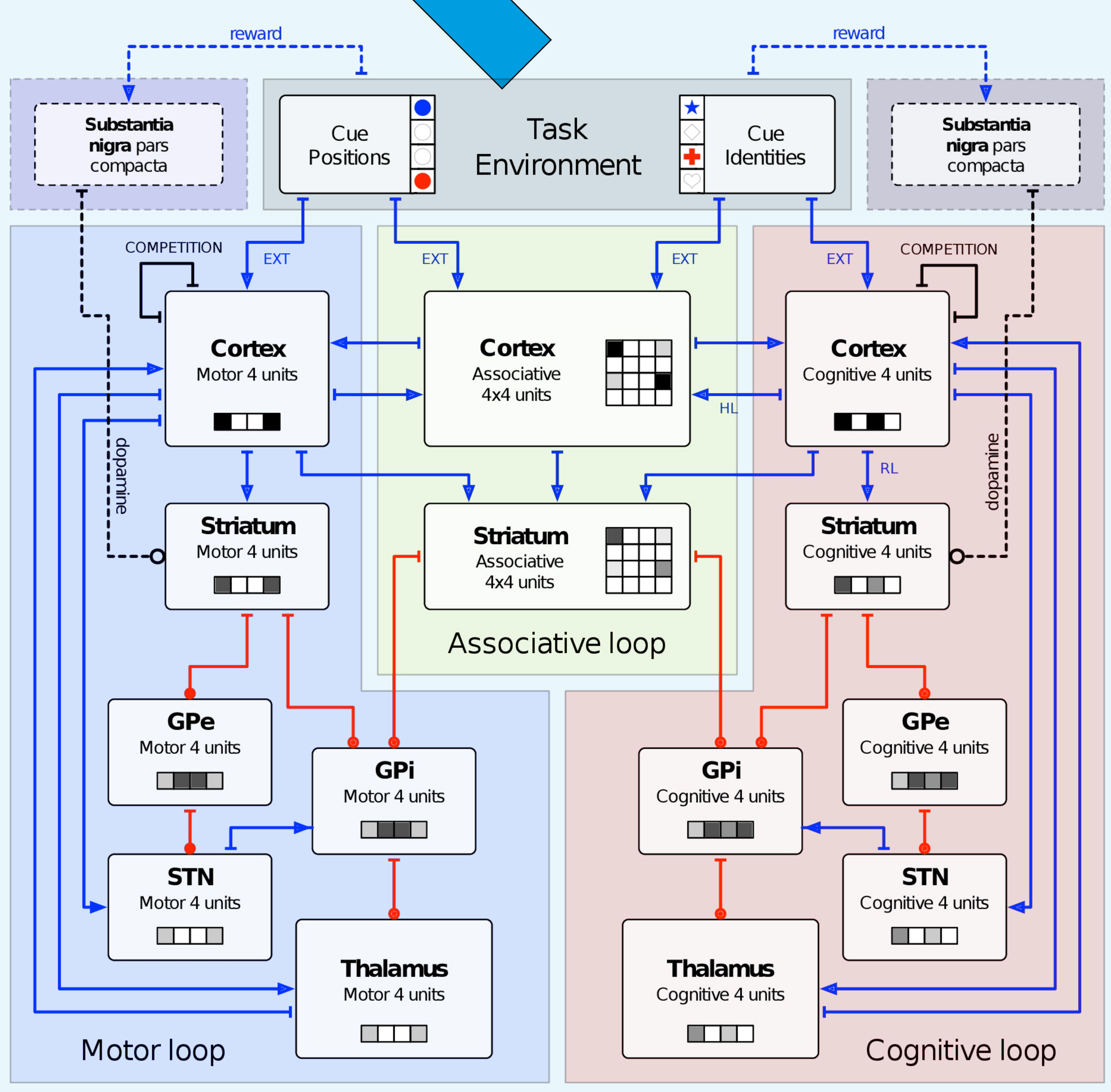


Meropi Topalidou, Daisuke Kase, Thomas Boraud, Nicolas P. Rougier

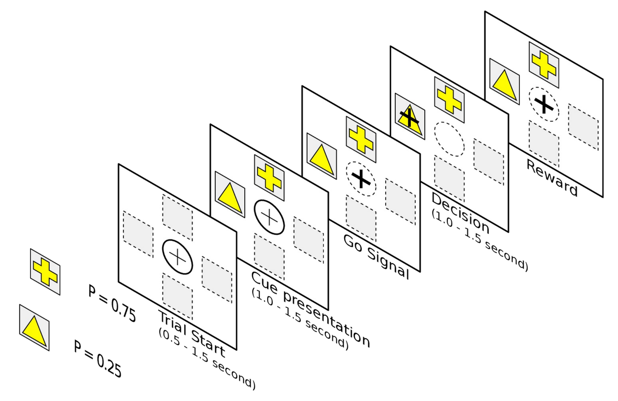
Abstract Even though there was early interest in the neural basis of automatic behavior in the beginning of the previous century, it is still unclear how we obtain and express habits. The dominant view of the 20th century was that novel behaviors are generated by cortex, because they require attention and flexible thinking. Contrary, habits are automatic, so they are primarily mediated by subcortical structures, more precise basal ganglia (BG). Nowadays, there is arising bibliography that argues exactly the opposite position. So the essential question is "Who's the teacher and who's the pupil?". Though our computational model of cortical-basal-thalamic closed loop, we show how **learning** has been **transferred from** the **basal ganglia to** the **cortex**, simply as a consequence of the statistics of the choice. Our theory was also tested on monkeys, which confirmed our predicted results.







Task For each trial, two cues are presented. After the choice of a shape, reward is delivered proportional to its fixed reward probability. The cue symbols and probabilities remained the same during a session.

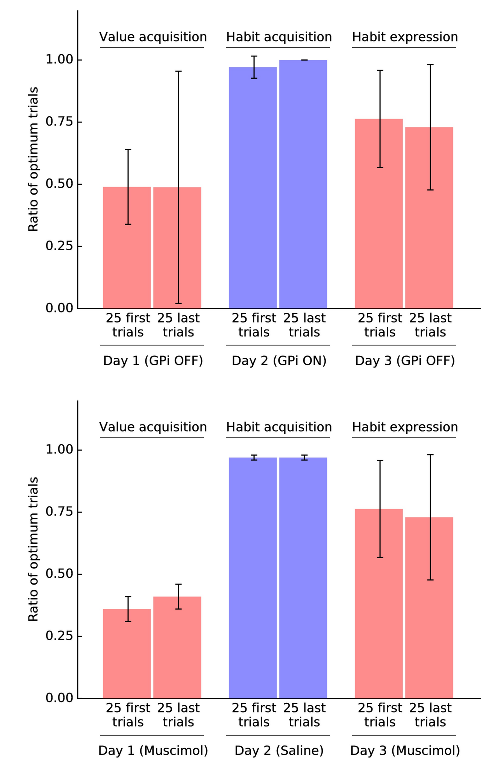


Experiments

Model In each simulation (out of 120), the model was tested first for 100 trials with inactive the main output of BG, GPi (D1), using the same set of stimuli. After-wards, it was tested for 120 more trials with active GPi (D2) this time, using the same set of stimuli as previously. FInally, the GPi was suppressed (D3) once more and the model was tested on the same stimuli for 120 trials.

Monkeys Two female macaque monkeys were tested on the task. 15 minutes before working session, a solution was injected bilaterally in GPi: either saline in saline condition (SC) or GABA agonist muscimol in muscimol condition (MC). On day 1, the two monkeys were trained for 6 and 7 sessions respectively in MC; using the same set of stimuli for each session. On day 2, the same number of sessions were conducted but in SC, using the same set of stimuli as in day 1.

Results



Model Upper figure shows that when the GPi output is suppressed (D1), the performance of the model is random with a 50% probability of choosing the best cue through session(chance level). However, after GPi inhibition has been removed, the model instantly reaches a very high performance level (D2). In the last test with suppressed GPi again (D3), the performances remain much higher than chance level.

Monkeys The lower figure shows that animals are unable to choose the best stimulus on D1 (in MC), however it shows a drastic change on behavior on D2 (in SC) when animals starts with near-optimal performance on the first 25 trials. The monkeys haven't tested to the third condition of the protocol (D3), but it is predicted their performanes to stay in a high level.

Model

The **model** is based on [1,2] and organized along three segregated cortico-basal-thalamic close loops (motor, associative and cognitive). A competition mechanism within each cortical group has been ad-ded [3], by using short range excitation and long range inhibitions, ensuring a unique cognitive and motor decision.

Learning occurs between the cognitive cortex and striatum through dopamine which modulates learning using reinforcement learning (RL; LTD & LTP). Hebbian learning (HL; LTP) at cortical level has also been added between cognitive and associative cortex. This learning is enforced once per trial, when a move has been made and independently of the actual reward.

Python code

[1] M. Guthrie, A. Leblois, A. Garenne, and T. Boraud. *Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study.* Journal of Neurophysiology, 109:3025–3040, 2013.
[2] Leblois A., Boraud T., Meissner W., Bergmann H., Hansel D. *Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia.* J Neurosci 26: 3567–3583, 2006.
[3] Piron C., Daisuke K., Topalidou M., Goillandeau M., N'guyen T., Orignac H., Rougier N.P., Boraud T. *The Globus Pallidus Pars Interna in Goal-Oriented and Routine Behaviors: Resolving a Long-Standing Paradox.* Accepted in Movement Disorders.

Membrane Potential : $\tau \frac{dV}{dt} = -V + I_{ext} + I_{syn} - T, \qquad U = f(V)$

Syhnaptic Input : $I_{syn}^{A\to B} = \sum_A gain \times G_B^A \times m_A$

Hebbian learning (I) : $\Delta W_{A\to B} = U_A \times U_B \times (W_{A\to B} - W_{min}) \times (W_{max} - W_{A\to B})$

Reinforcement learning $\Delta W_{A\to B} = \alpha \times PE \times U_B$
$PE = Reward - V_i$

Modelling — Meropi Topalidou
Experiments — Daisuke Kase, PhD
Experiments — Thomas Boraud, MD, PhD
Modelling — Nicolas Rougier, PhD